

## Learning Analytics with Big Data: A Framework

**Rafi Ahmad Khan**

Presently at Faculty of Computer and Information System, Islamic University of Madinah, Madinah, Kingdom of Saudi Arabia  
& Department of Management Studies, University of Kashmir, India

*Corresponding Author: mca\_rafi@yahoo.com*

DOI: <https://doi.org/10.26438/ijcse/v7i4.349353> | Available online at: [www.ijcseonline.org](http://www.ijcseonline.org)

Accepted: 12/Apr/2019, Published: 30/Apr/2019

**Abstract**—Learning Management Systems (LMS) introduced in 1990's support online learner-centered model by complete management of teaching, learning and assessment. These LMS help teachers to deliver course related material to their students, manage quizzes, exams and other course related tasks, evaluate student performance and manage other activities. The LMS tools like Moodle, Blackboard, Sakai, EvalTool, Dokeos etc., have produced outstanding results for both teachers as well as students. Since most of students use laptops and smart phones to access LMS for their learning related activities, therefore, their online activities generate huge volume of data that educational institutions can use to improve the performance of their teaching and learning activities. This huge volume of data can give a deep insight into the teaching and learning activities if analyzed properly. Learning Analytics(LA) is a new emerging field that analyzes this big data and develop the models that can predict the performance of student, detect the students who are most likely to drop out from the courses, prepare the reports, provide the intelligent and instant feedback to students, recommend the courses to students based on their interests and assesses the skills developed by the students. Therefore, by using learning analytics, teachers, students, faculty, and administrators can develop more engaged and effective teaching and learning techniques. Keeping in view the importance of LA, this paper discusses the LMS, LA and a framework for using LA on the LMS data.

**Keywords**— Learning Analytics (BA); Data Mining; Dashboard; Scorecard; Learning Management Systems (LMS)

### I. INTRODUCTION

That fast development of information and communications technologies (ICT) during last two decades has made impact on all fields of life including education. Due to emergence of ICT, the traditional face-to-face teaching and learning has transformed as a result of introduction of online resources and material. Now a large number of educational institutions around the world are using the internet in teaching and learning, especially Learning Management Systems (LMSs), also known as Virtual Learning Environments (VLEs) [1]. LMSs enhances the student learning by delivering the content and resources online, facilitating the online quizzes, presentations and assignments, and by providing the online discussion forums [2]. Moreover, LMSs allow teachers to provide and manage these resources in a relatively easy and integrated way.

As LMS monitors and stores every action, the analysis of such data can give better understanding of online behavior of students, resulting in enhanced teaching and learning. This analysis of LMS data is called Learning Analytics [3] and is defined as a means to help educators observe,

understand, and support students study behaviors ' and change their learning environments [4].

Keeping in view the importance of LA, this paper discusses the concept of LA, its uses in educational institutions and has also given general framework of LA that describes how it uses the big data generated by LMS.

### II. LEARNING ANALYTICS(LA)

Global developments in higher education show a shift from the traditional teacher-centered model that stresses on lectures, tests and other methods of assessments towards the technology supported learning that emphasize on what students know and can do. Learning Management Systems (LMS) are the web-based applications that facilitate the capture and storage of all course activities like, quizzes, exams, assignments, projects etc. along with the scores and evaluation of each of these activities. LMS work as an automated system to gather data and convert this data into text files and visual data [5].

LMSs are valuable tool for developing collaborative environment among learners by using active learning techniques, time on task, giving quick feedback, high

interactive abilities, and providing different ways of learning [6]. LMSs software's like Sakai, Blackboard, EvalTool, TalentLMS, Moodle etc. have made huge impression on teaching, learning and assessment by providing online distribution of content, schedules, safe access, collaboration tools, chat, homework, quizzes, etc. The use of these tools by the educational institutions result in generation of huge volume of data that can be used to evaluate and predict the performance of students hence results in overall improvement of teaching and learning.

But most of this huge data which mostly remains unused [7]. This big data can be analyzed to get the insight into the academic performance of teaching and learning [8]. The analysis of LMS data is often stated as Learning Analytics (LA) [3]. Learning Analytics (LA) collects, measures, analyzes data and generates reports of this data about students in order to understand and improve learning and its environment [9]. LA is the interpretation of the students data generated and collected in order to assess and evaluate their academic progress, forecast their future performance, and identify the potential issues [10]. Learning analytics (LA) is a rapidly evolving research discipline that aims at "measurement, collection, analysis, and reporting of data about learners and their contexts, for purposes of understanding and optimizing learning and the environments in which it occurs" [11]. LA is closely related to other fields of study including business analytics, educational data mining, academic analytics, web analytics, and action analytics.

Hamelin & Workman identified some of the uses of LA in educational institutions as [12]:

- It can identify those students who are at risk of leaving a course /program and offer the helpful interventions intended to improve retention
- It can offer recommendations to students related to course material and learning activities.
- It can identify the need for educational improvement and can also measure the results of improvements.
- It can recommend the tailored course offerings.
- It can Identify the well performing teachers as well as those teachers who need assistance with their teaching methods.
- It can help in the student recruitment process

Considerable study in the area of Learning Analytics has used LMS data to develop models that can predict the

performance of students and to predict the students who are most likely to drop out of a course (Attrition Risk) [13] [14] [15].

Drachsler & Greller has given the framework of LA as shown in fig 1 [16]. It describes the various dimension of

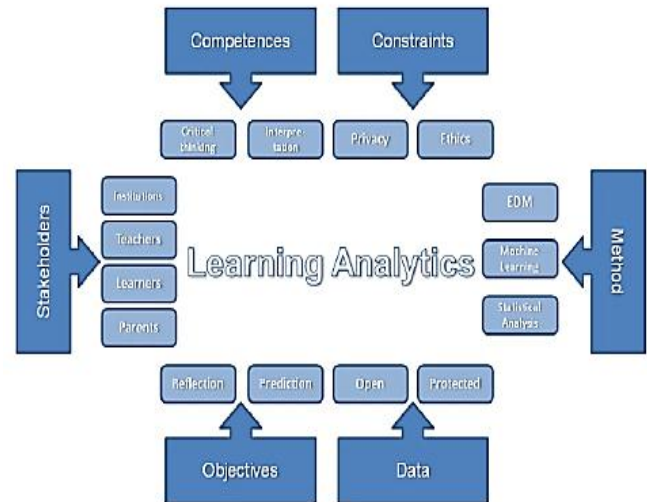


Fig 1: Framework of Learning Analytics [16]

Learning Analytics as described below:

**Stakeholders:** Stakeholders include teachers, academic administrators, students, and education policy makers.

**Objectives:** These are the goals that education institutions aim to achieve.

**Data:** Various types of data and datasets that are generated by LMS in teaching and learning environment. It can be related to teaching, learning, assessment and evaluation.

**Methods:** They include the procedures, techniques, and models employed to analyze the data related to education.

**Constraints:** These are a limitation or restriction.

**Competences:** These are abilities to do a particular activity to a prescribed standard.

### III. FRAMEWORK OF LEARNING ANALYTICS (LA)

A general framework of LA consists of three layers viz., Data layer, Analytics layer, Reporting/ Visualization Layer, Access layer. These layers are discussed below:

**(i)Data layer:** Data layer is concerned with acquisition/capture of data from various sources. Data from LMS as well as other sources like Academic Management Systems, Social Media etc. is acquired. Since data is collected from dissimilar sources in raw format which is not feasible for the analysis, therefore, it needs to be transformed into a right format before it is used by a particular LA method [17]. This phase is important for the successful discovery of valuable and hidden patterns from the data. Data preprocessing prepares the data for analytics. The preprocessed data is stored in a data repository before analyzing it.

**(ii)Analytics Layer:** In this layer, data from previous layer is analyzed by using Descriptive/Predictive/Prescriptive analytics as discussed below:

#### Descriptive Analytics – What happened

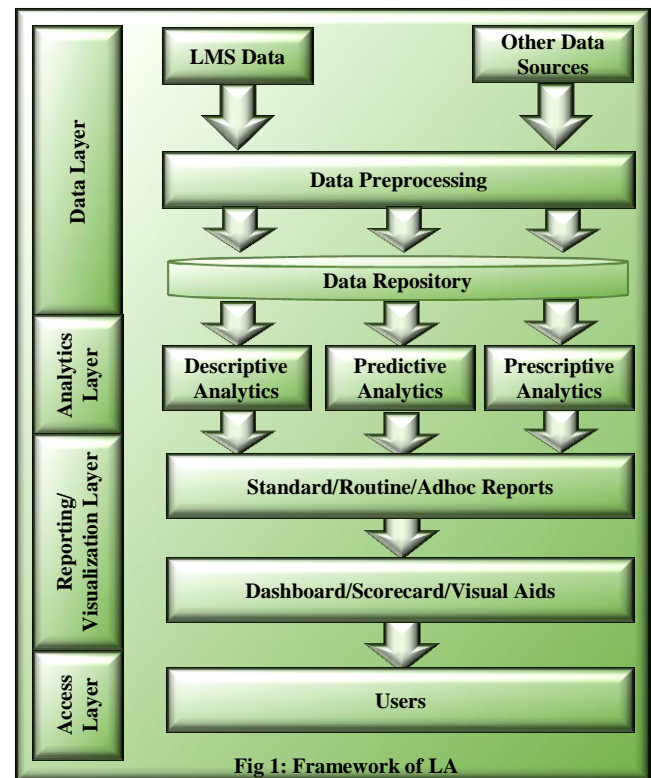
Descriptive analytics gives insight into the past trends and performance. It provides insight from historical data with reporting, scorecards, clustering etc. [18] which gives understanding of student's behavior. Descriptive analytics provides routine, regular and Ad hoc reports that can help the education institutions to look at the facts like, what has happened, where, and how often. It can execute specific queries so that teachers and administrators can examine the exact problem. Visualization has become an important component of descriptive analysis as it can develop powerful insights into the actions and their effect on teaching and learning.

Examples of descriptive analytics may be as simple as how many students took a course, the grades students got in a quiz, how many times they accessed LMS and which functions of LMS they used. These predictive analytics functions are provided in the LMS.

#### Predictive Analytics – What will happen

Predictive analytics concentrates on predicting the unknown future events based on historical data. It tries to answer the question "What might happen?" and is a step ahead of descriptive analytics. According to Raden [19], the objective of predictive models is to understand the causes and relationship in the data so as to make accurate predictions. It includes a wide variety of methods and techniques including modeling, machine learning, and data mining to analyze present and past facts to make predictions about future events [20].

In LA, it is used to predict the students who can be underperforming in a given course in the exam which can be helpful to teachers to improve their teaching and course material. It can also identify the students that are at risk of dropping out of the course or program and interventions can be made to assist such students in completing their course or program. Predictive analysis has enabled the development of such an alert system in education. It can be used to enhance



the student's learning process by providing real-time feedback, or by enriching the learning experience.

#### Prescriptive Analytics – Prescribe Solutions for Various Outcomes

Prescriptive Analytics allows users to "prescribe" a number of course of action for a given situation and guides them towards a solution. It goes beyond descriptive and predictive analytics by prescribing one or more possible courses of action [21].

In LA, it is used to study the past actions of students to deliver the evidence-based information on personalized actions and resources that can help students in enhancing their understanding and performance [22].

Based on the student data collected from the LMS, it can create suggestions on various instructional resources and tools to use so that students can make greater impact.

**Reporting/ Visualization Layer:** Various tools used in this layer are as:

**Dashboards:** They are the tools for visualization of important business data presented in the form of graphic indicators, charts and tables [23], [24]. A digital dashboard provides the user a graphical high-level view of business processes that can be drilled-down to find more detail on a particular business process. This level of detail is often buried deep in the enterprise's data, making it otherwise

concealed to a business manager [25]. It gives users ability to instantly respond to information being presented and also provides the facility of drill down and root cause analysis of situations at hand [26]. It organizes and reflects the information into the user interface in an easy, interactive and intuitive manner. It can facilitate users to generate their individual dashboards with various features like tables, reports, pivot tables, graphs, and prompts.

**Balanced Scorecards:** The Balanced Scorecards (BSC) concept was presented by Kaplan [27]. They are semi-standard structured report, supported by design tools and techniques, that can be used by managers to keep track of the execution of activities by the staff within their control and to monitor the consequences arising from these actions [28]. They are performance metric to identify and improve various internal functions and their resulting external outcomes. BSC enables organizations to put the strategy into practice by measuring and delivering feedback to the organizations.

**Reports:** These are the written documents relating to the situation and can be created by the end users by supplying parameter data. The pre-executed report results are cached in order to support interactive and high-performance viewing of these reports.

**Ad hoc Reports:** Contrary to standard reports which are predefined and routinely processed, ad hoc reports are generated when the need arises. They enable users to produce their own customized reports without relying on IT team.

**Alert:** is a type of reports that is automatically triggered when an event occurs, e.g. an e-mail or SMS message is sent to the students/teacher when the performance of the student is below average.

#### IV. TECHNIQUES

Daud et. al., has given the overview of various data mining techniques for Advanced Learning Analytics that can be used to analyze the data generated in educational institutions especially universities [29]. It includes Neural Networks, Support Vector Machines (SVM), Regression, Feature Selection, Decision Trees, Prediction Models etc. that can be used in Learning Analytics.

In predictive analytics, various types of regression techniques including Linear, Discrete, Logistic, timeseries models, Classification and regression trees (CART) etc. are used. Machine learning techniques such as Neural networks, Multilayer perceptron (MLP), Radial basis functions, Support vector machines, Naïve Bayes, k-nearest neighbors etc. are used.

#### V. CONCLUSION

Most of these institutions world over are now using the LMS to manage the teaching and learning efficiently and effectively. With the use of these tools, they are flooded with academic data of students related to their internal assessments, quizzes, midterms, lab, finals as well as data related to teaching including instructional material and resources. The educational institutions can use this LMS generated data by processing it and converting it into valuable information/knowledge that can be used to increase the performance of students as well as help them in decreasing the rate of attrition of students in different courses/ programs. This will help them to increase the efficiency and effectiveness of both teaching and learning in their institutions. LA has emerged as the most widely used technique in educational institutions for evaluation of the LMS data to convert it into information and knowledge in the form of reports, tables, visualization and other analytical tools which gives these institutions an insight in to the performance of students as well as teaching and learning. The framework given in this paper gives an outline of the LA and how it can be implemented by using the big data generated by the LMS.

#### ACKNOWLEDGEMENTS

This work is part of the research study titled “Use of Learning Analytics on Big Data to Predict the Performance and Attrition Risk of Students in Islamic University of Medina” supported by Deanship of Research, Islamic University of Madinah, Madinah, Kingdom of Saudi Arabia.

#### REFERENCES

- [1] S. Retalis, A. Papasalouros, Y. Psaromiligkos, S. Siscos and T. Kargidis, “Towards networked learning analytics—A concept and a tool,” in *5th Int. Conference on Networked Learning*, Lancaster, 2006.
- [2] A. A. Pina, “An overview of learning management systems,” in *Virtual Learning Environments: Concepts, Methodologies, Tools and Applications, 1st ed*, Louisville, KY, USA, Information Resources Management Association (USA), 2012, p. 33–51.
- [3] G. Siemens and R. S. Baker, “Learning analytics and educational data mining: Towards communication and collaboration,” in *2nd International Conference on Learning Analytics Knowledge*, 2012.
- [4] A. Rubel and K. Jones, “Student privacy in learning analytics: An information ethics perspective,” *The Information Society*, vol. 32, no. 2, pp. 143-159, 2016.
- [5] S. Wallden and E. Makinen, “Educational data mining and problem-based learning,” *Informatics in Education*, vol. 13, no. 1, pp. 141-156, 2014.
- [6] J. Wang, W. J. Doll, X. Deng, K. Park, M. Ga and M. Yang, “The impact of faculty perceived reconfigurability of learning management systems on effective teaching practices,” *Computers*

- & Education, vol. 61, p. 146–157, 2013.
- [7] K. Parimala, G. Rajkumar, S. Ruba and S. Vijayalakshmi, “Challenges and Opportunities with Big Data,” *International Journal of Scientific Research in Computer Science and Engineering*, vol. 5, no. 5, pp. 16-20, 2017.
- [8] S. Shirasath, V. A. Desale and A. D. Potgantwar, “Big Data Analytical Architecture for Real-Time Applications,” *International Journal of Scientific Research in Network Security and Communication*, vol. 5, no. 4, pp. 1-8, 2017.
- [9] P. Long, G. Siemens, G. Conole and D. Gasevic, “Announcing open course: Learning and knowledge analytics,” in *1st Int. Conf. Learning Analytics and Knowledge*, Banff, AB, Canada, 2011.
- [10] L. Johnson, R. Smith, H. Willis, A. Levine and K. Haywood, The 2011 Horizon Report, Austin, Texas: The New Media Consortium, 2011.
- [11] G. Siemens, “Learning Analytics: The Emergence of a Discipline,” *American Behavioral Scientist*, vol. 57, no. 10, p. 1380–1400, 2013.
- [12] M. V. Harmelen and D. Workman, “Analytics for Learning and Teaching,” *Centre for Educational Technology & Interoperability Standards*, vol. 1, no. 3, 2012.
- [13] R. S. J. Baker and K. Yacef, “The state of educational data mining in 2009: A review and future visions,” *Journal of Educational Data Mining*, vol. 1, no. 1, p. 3–17, 2009.
- [14] C. Romero and S. Ventura, “Educational data mining: A review of the state of the art,” *IEEE Transactions on Systems, Man, and Cybernetics, Part C*, vol. 40, no. 6, p. 601–618, 2010.
- [15] S. B. Shum and R. Ferguson, “Social learning analytics,” *Journal of Educational Technology & Society*, vol. 15, no. 3, p. 3–26, 2012.
- [16] H. Drachsler and W. Greller, “The Pulse of Learning Analytics Understandings and Expectations from the Stakeholders,” in *Learning Analytics and Knowledge*, 2011.
- [17] N. S. Gill, “Data Preparation, Preprocessing and Wrangling in Deep Learning,” 23 May 2017. [Online]. Available: <https://www.xenonstack.com/blog/preparation-wrangling-machine-learning-deep/>. [Accessed 24 March 2019].
- [18] R. Saxena and A. Srinivasan, *Business Analytics-A Practitioner’s Guide*, Springer, 2013.
- [19] N. Raden, “The Foundations of Analytics: Visualization, Interactivity and Utility. The ten principles of Enterprise Analytics,” Spotfire, Inc, Somerville, U.S.A, 2010.
- [20] ECAR, “The Predictive Learning Analytics Revolution: Leveraging Learning Data for Student Success,” ECAR-ANALYTICS Working Group, Louisville, CO: ECAR, 2015.
- [21] Halo, “Descriptive, Predictive, and Prescriptive Analytics Explained,” 24 March 2019. [Online]. Available: <https://halobi.com/blog/descriptive-predictive-and-prescriptive-analytics-explained/>.
- [22] L. O’Farrell, “Using Learning Analytics to Support the Enhancement of Teaching and Learning in Higher Education,” National Forum for the Enhancement of Teaching and Learning in Higher Education, Dublin 2, Ireland, 2017.
- [23] A. Kirtland, “Executive Dashboards,” 20 1 2006. [Online]. Available: <http://dssresources.com/papers/features/kirtland/kirtland01202006.html>.
- [24] A. Vasiliu, “Dashboards and scorecards: Linking management reporting to execution,” 30 4 2006. [Online]. Available: <http://dssresources.com/papers/features/vasiliu/vasiliu04302006.html>.
- [25] Gravic, “The Evolution of Real-Time Business Intelligence and How To Achieve It Using HPE Shadowbase Software,” Gravic, Inc., Malvern, PA, USA, 2017.
- [26] D. Chappelle, “Big Data & Analytics Reference Architecture,” Oracle Corporation, CA, USA, 2013.
- [27] R. Kaplan and D. Norton, *The balances scorecard*, Boston: Harvard Business Press, 1996.
- [28] C. M. Olszak, “An overview of information tools and technologies for competitive intelligence building: Theoretical approach,” *Issues in Informing Science and Information Technology*, vol. 11, pp. 139-153, 2014.
- [29] A. Daud, N. R. Aljohani, R. A. Abbasi, M. D. Lytras, F. Abbas and J. S. Alowibdi, “A Predicting Student Performance using Advanced Learning Analytics,” in *Proceedings of the 26th International Conference on World Wide Web Companion*, Perth, Australia, 2017.